Measuring vowel percepts in human listeners with behavioural response-triggered averaging

W. Owen Brimijoin (MRC Inst. of Hearing Res., Glasgow, United Kingdom), Emily Tilbury (School of Eng., Univ. of Glasgow, New Acoust., 1 Aurora Ave., Queens Quay, Clydebank G81 1BF, United Kingdom, emily@newacoustics.co.uk), Michael A. Akeroyd (MRC Inst. of Hearing Res., Glasgow, United Kingdom), and Bernd Porr (School of Eng., Univ. of Glasgow, Glasgow, United Kingdom)

A vowel can be largely defined by the frequencies of its first two formants, but the absolute frequencies for a given vowel vary from talker to talker and utterance to utterance. Given this variability, it is unclear what criteria listeners use to identify vowels. To estimate the vowel features for which people listen, we adapted a noise-based reverse-correlation method from auditory neurophysiological studies and vision research (Gold et al., 1999). Listeners presented with the stimulus, which had a random spectrum with levels in 60 frequency bins changing every 0.5 s, were asked to press a key whenever they heard the vowels [a] or [i:]. Reverse-correlation was used to average the spectrum of the noise prior to each key press, thus estimating the features of the vowels for which the participants were listening. The formant frequencies of these reverse-correlated vowels were similar to those of their respective whispered vowels. The success of this response-triggered technique suggests that it may prove useful for estimating other internal representations, including perceptual phenomena like tinnitus.

References:

Gold, J., Bennett, P. J., and Sekuler, A. B. (1999). “Identification of band-pass filtered faces and letters by human and ideal observers,” Vis. Res. 39(21), 3537–3560.